

DOCUMENT RESUME

ED 366 587

SP 035 013

AUTHOR Siegel, Donna F.; Hanson, Ralph A.
TITLE A Longitudinal Evaluation Model for Teacher Education Programs.
PUB DATE Feb 94
NOTE 64p.; Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education (Chicago, IL, February 16-19, 1994).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)
EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Elementary Secondary Education; *Evaluation Criteria; Evaluation Methods; Evaluation Research; Followup Studies; *Formative Evaluation; Higher Education; Longitudinal Studies; Models; Operations Research; Program Costs; *Program Evaluation; *Summative Evaluation; *Teacher Competencies; Teacher Competency Testing; *Teacher Education Programs
IDENTIFIERS *Confirmation (Strategy)

ABSTRACT

This presentation outlines, in operational terms, the components and products of a longitudinal evaluation model for teacher education programs, and discusses an approach to program evaluation that embodies activities within three taxonomical phases referred to as formative, summative, and confirmative evaluation. When applied to competency-based certification, a longitudinal evaluation model for a teacher education program is created that can link procedures and costs to actual effects and competencies. The paper describes this approach, the products produced, and the criteria for completion for each of the three evaluation phases. Tasks outlined set the agenda for a complete teacher education program evaluation system that would track policies and students in each specialty area program during its operational life. In this way, firm estimates of a program's long-term effects, benefits, and costs can be derived. Appendices provide criteria for completion of a formative, summative, and confirmative evaluation; examples of professional teaching competencies and indicators of each competency; information for possible consideration in a longitudinal student data base; and examples of items to include in the entrance, exit, and follow-up assessment instruments. (Contains 27 references.) (LL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

A Longitudinal Evaluation Model for Teacher Education Programs¹

Donna F. Siegel
The University of Science and Arts of Oklahoma
P.O. Box 82345
Chickasha, Ok 73018-1000

Ralph A. Hanson
Hanson Research Systems
12889 Knott St., Suite 219
Garden Grove, CA 92641-3925

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL
IN OTHER THAN PAPER COPY HAS BEEN
GRANTED BY

D. Siegel

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

¹Paper presented at the annual meeting of the American Association of College Teachers of Education, Chicago, IL, February 13-16, 1994.

BEST COPY AVAILABLE

ABSTRACT

This paper presents a discussion of a longitudinal approach to program evaluation that embodies activities within three taxonomical phases referred to as: formative, summative, and confirmative evaluation (Siegel and Hanson, 1991). When applied to a competency-based, certification model it creates a longitudinal evaluation model for a teacher education program that can link procedures and costs to actual effects and competencies. This paper describes this approach, the products produced, and the criteria for completion for each of the three evaluation phases for a teacher education program. The tasks outlined in this paper set the agenda for a complete teacher education program evaluation system that would track policies and students in each specialty area program during its operational life. In this way, firm estimates of a program's long-term effects/benefits and costs can be derived.

A Longitudinal Evaluation Model
for Teacher Education Programs

D.Siegel and R. Hanson

Every year, university, state department, regents, and accreditation groups (e.g., NCATE), among others, request a variety of information regarding teacher education programs (Haywood, 1989; Cooley, 1983). They want to know, for example:

- Is a particular program producing the desired or anticipated outcomes?
- What specific kinds of educational experiences and configuration of courses are required to develop certifiable and competent teachers and what are the costs associated with this process?
- Is a four- or five-year teacher preparation program better?

Moreover, escalating costs, tight budgets, and a general trend toward greater accountability provide incentives for such groups to seek this kind of information. These policy groups, as well as the general public, now question the value and worth of specific teacher education programs (e.g., Pearce, 1991; Kelly, 1994). If teacher education institutions

cannot provide evidence to the contrary, then these constituencies are at liberty to impose any of the latest educational fads or snake oil remedies that come along - and they do! For example, in many states new waves of reforms for teacher education programs are now being implemented without even having determined the results of those implemented only a few years previously. Examples of the current crop of "reforms" include:

- Implementation of new requirements for both traditional and alternative certification;
- Elimination of specialty area programs or courses of instruction that are perceived as not cost-beneficial;
- Requiring additional course work and/or other significant changes in graduation requirements and program policies;
- Major reconfigurations of either specialty area or entire Teacher Education programs because of perceived "better" ways of developing competent teachers; and/or
- Placing programs on reduced budgets and with fewer resources.

While many of these "reforms" may be appropriate, under current conditions it will be very difficult to determine this with any certainty. That is, though such changes are carried out with the best intentions, these and other educational reforms are typically made with little knowledge of how they will impact Teacher Education programs over a period of time. These decisions are based only on speculative (and often very questionable) reasoning.

As a result, it is now imperative that professionals associated with teacher education programs across the country have the data to provide defensible answers to some very reasonable questions; both to justify the policies and practices of our nation's teacher preparation programs and to avoid capricious program changes. In essence, teacher education programs, like other publicly supported programs, must demonstrate their worth in order to continue to receive the tax-payer's support.

Traditionally, both administrators and politicians have sought to provide expedient responses to the variety of requests and mandates they receive each year in regard to teacher education programs. The bad news

is that this approach has typically resulted in more controversy than problem resolution. The good news is that an empirical means of justifying a program's architecture, policies, and practices does exist; it simply requires use of different methods and procedures to develop and implement such an approach (Provus, 1971; Cooley, 1980;1983; Nichols, 1989).

This paper presents a discussion of one approach that can provide the needed information. This longitudinal approach to program evaluation embodies activities within three taxonomical phases referred to as: formative, summative, and confirmative evaluation (Siegel and Hanson, 1991). When applied to a competency-based, certification model it creates a longitudinal evaluation model of a teacher education program that can link procedures and costs to actual effects and competencies. This paper describes this approach, the products produced, and the criteria for completion for each of the three evaluation phases for a teacher education program. In addition, examples of the kinds of information that can be generated by the model are provided.

Longitudinal Evaluation Model Background

Teacher Education Programs historically have been viewed in terms of a competency-based, certification model (e.g., Kinney, 1964). That is, today as in the past, the primary goal of such programs is to provide both a screening device to "protect" the public and a means for individuals to obtain competency, licensure, and certification in their career specialty areas. Furthermore, teacher certification, like certification in other professions, is linked to accreditation agencies and professional associations and usually has some impact on an individual's wages and salary.

Thus, Teacher Education program are designed primarily for the purpose of preparing individuals to gain the knowledge and skills necessary to be certified as teachers. The problem with this model is that, although literally hundreds of skills and behaviors have been linked to effective teaching, the exact professional and educational process remains to be defined and validated. That is, out of all of the skills and behaviors identified as related to teaching competency, both analytically and empirically, researchers are still unable to:

1. identify and agree upon a set of essential teacher competencies that must be attained in order for individuals to become successful classroom teachers; and
2. measure and monitor these competencies in a teacher's professional life to empirically validate that model.

The ability to do this, however, is rapidly becoming a requirement for managing credible and effective teacher education programs (Raschke, 1993). In the recent past, for example, allowing incompetent teacher education program graduates to become certified was considered to be of little consequence; there was rarely any national public outcry regarding incompetent classroom teachers. The general consensus appeared to be that, overall, a few poor teachers in our nation's classrooms could do little harm to the children involved. This notion has changed considerably; in most states, teaching is now becoming regarded as a high stakes certification area.

Until recently, however, neither the methodology nor the technology has existed to allow researchers to empirically link the particular higher educational

experiences of teachers to their success and effectiveness in the classroom. To accomplish this task required a combination of advanced technology and the methodology for a longitudinal approach to program evaluation.

Today both of these requirements have been met. The advances made in computer technology and the methods and procedures developed for a three-phase program evaluation model (Siegel and Hanson, 1991) now provide the means to empirically validate a certification model as well as the capability to compile, monitor, and report the program effects. Thus, when the framework of the three-phase program evaluation model is applied specifically to teacher certification, a longitudinal evaluation model for Teacher Education programs emerges.

Central to this model is the notion that a program is a dynamic system with a useful life (e.g., Hanson, 1986; Raizen and Rossi, 1981; Jordan and Machesky, 1990). That is, to be valid, a program must first meet the specific technical criteria that operationalize it (i.e., through formative evaluation). Once this has been done, the program must be continually monitored

and managed during its implementation (i.e., through summative evaluation). Next, the immediate and enduring costs and effects must be determined as the students become certified career teachers (i.e., through confirmative evaluation). The results of these evaluations are then fed back into the program review process. Thus, the full evaluation process continues to operate until such a time that the program is either no longer needed, practical, or politically/educationally in vogue.

Although the exact length of time required to fully implement this model will vary, Figure 1 provides a suggested time frame for each phase. What is

Insert Figure 1 About Here

important to note in this regard is that the phases are taxonomical. Therefore to be useful, the first phase, formative evaluation, is critical for both the summative and confirmative phases of the evaluation. Hence, it is essential to take extra time and effort in each phase to carefully ensure that all of the criteria for completion have been met.

Phase I: Formative Evaluation

The main purpose of the formative evaluation is to operationalize all teacher education specialty area programs and establish their content validity. This means that a technical report for each specialty area program should document that the program has met all of the criteria for the completing a formative evaluation. The complete list of these criteria are presented in Appendix A. Collectively, the products produced by meeting these criteria help to ensure the quality and stability of each specialty area program within a Teacher Education program over time.

Traditionally, as indicated in Figure 2, explanations of formative evaluation refer only to those procedures that take place during a program's development to improve specific aspects before it is widely distributed and implemented (e.g., Theobald, 1985; Dignan, 1986). However, because most teacher

Insert Figure 2 About Here

education programs have already been developed and are in use, these procedures and products must be reconfigured and further developed in order to meet the

criteria for completing the formative evaluation phase. Producing the program products to meet these criteria is expected to take place over a three- to four-year period of time (See Nichols, 1989 for a more detailed account of issues associated with validating program effectiveness and implementing outcomes assessments). These products, which are listed and described in the following section, must be developed for each specialty area program by the teacher education administration and faculty involved in each program. Fortunately for these individuals, and particularly for those who have been involved in the NCATE accreditation process, many of these products are familiar and, in many cases, may be either already developed or in the process of development.

Formative Evaluation Products to be Developed

I. A Preliminary set of Teacher Competencies for Each Program Specialty Area

This is essentially a report containing the institution's teacher certification model described in terms of a specified set of teaching behaviors and course outcomes. Ideally, all state departments of

education should have a certification model consisting of the set of teaching competencies required for state licensure and certification in each teaching area. Each institution's teacher certification model should then consist of a set of competencies which, at the very least, would be fully sufficient to meet the state teaching certificate requirements. Furthermore, these competencies would be linked to each specialty area program's courses of instruction, examinations, and other requirements (including practice teaching) that comprise the curriculum.

Collectively, these experiences provide instruction and training in all model competency areas within each specialty program. Theoretically, at least, completing this series of curriculum requirements will enable students to meet state teacher certification requirements and become certified teachers with the necessary competencies in their specific teaching field. Conversely, by carefully analyzing the objectives, requirements, and outcomes for each course of instruction within a specialty area program, the specific, measurable competency areas that comprise the model will be operationally defined.

At this point, it is important to differentiate among competencies, outcomes, and objectives. Competencies are defined as completely developed skills, usually higher-level and generalizable across contexts. They are the result of instruction and training that has taken place over a considerable period of time (e.g., a year or more). Competency attainment requires taking a series of course work, coupled with intern experiences (e.g., field experiences, student teaching) for over a period of at least a year or more. Broad skill areas such as classroom organization and management is one example of a teaching competency; Appendix B presents other examples, as well as some indicators of each competency.

Objectives, on the other hand, are more narrowly defined, involve fairly simple skills (e.g., learning a concept), and are considered to be lower-level, "on-route" skills and behaviors. Objectives may be either taught or learned within several hours of instruction (e.g., in one or two lessons). They are used primarily to develop the all of the resources, procedures, activities, and exams for a course of instruction.

Once all of these course materials are compiled and placed into operation, the objectives become of secondary interest; particularly after classroom experience indicates that the course is achieving the desired outcome(s).

Outcomes lie somewhere in between objectives and competencies. They are the aggregate product of a course of instruction and are derived from the course resources and procedures. While a typical course has a number of objectives that are linked to a set of resources and instruction, outcomes are the composite skills and knowledge that these objectives develop through the instructional process. Moreover, outcomes are much fewer in number (i.e., 1 to 3 per course) than objectives and are linked more closely to the overall instruction provided by, and the grade received for, a course. Outcomes are the formal assessment of a course of instruction.

Thus, viewed in terms of a typical teacher education program comprised of courses of instruction, a course might have a dozen or more objectives; however, the same course typically would not address more than three outcomes. These outcomes would, in

turn, provide the "higher-level" skills that contribute, in combination with other experiences, to the development of one or more teaching competencies.

Figure 3 provides an example of the relationship among these terms showing a hypothetical program with seven education courses and five competency areas. The specific kinds of education courses developed for a teacher education specialty area program are determined by a knowledge base and the admission criteria. As the model indicates, the program's knowledge base includes, among other things, all of the state certification requirements and other kinds of knowledge, skills, and behaviors identified as necessary for teacher competency by professional associations, research literature, professional boards, and institutional policies and values.

Insert Figure 3 About Here

Note that each course produces between one and three outcomes. Conceivably, as illustrated by the dotted lines in Figure 3, some outcomes may be developed only after completing two or three related courses. Since some courses are taken in a series (e.g., math, reading, field experiences), the defined

outcomes may apply only to those students who have completed the full series of courses. For example, two or three outcomes might be established for those students who have completed the set of courses that comprise all field experiences. Moreover, although the majority of these field experience course outcomes would probably tie into one or more of the identified competency areas, all of these course outcomes would not necessarily do so. However, as the figure indicates, it is the course outcomes that provide students with the necessary professional skills and behaviors for developing one or more of the program's teaching competencies.

Viewing teacher education programs in terms of teaching competencies represents a major change in focus from previous educational models. Traditional models typically place their focus on educational objectives and, more recently, outcomes - not competencies. In a certification model the focus is, as it should be, on a specified set of competencies. Thus, all instruction is developed and implemented with the notion of developing a specific set of competencies. Moreover, a certification model

legitimizes the inclusion of important characteristics of successful teaching that are less conducive to being taught, such as interpersonal competency, professionalism, and adaptability.

Identifying the teacher competencies to be included in the institutional model requires taking the following into account:

1. Competencies identified by state departments of education as being required for certification;
2. Competencies required by state public school districts to obtain tenure and remain certified;
3. Competencies identified by research and professional organizations as necessary for specialization areas;
4. Competencies required by state regents, legislative commissions, and professional standards boards;

Once the program's teacher competency areas are identified and defined, they are included in the program's technical report to: a) document the analytical framework for each specialty area program, b) provide justification for the program's courses of

instruction, and c) show evidence of its content validity.

II. Operational Assessments of Each Teacher Competency Area

After identifying and defining the teaching competency areas, the next task is to develop an operational assessment for each. Although developing these competency area assessments could feasibly be the responsibility of each institution's teacher education program, they most logically should be developed by a state's department of education and constitute the primary task in the design and development of a statewide certification and assessment model. This arrangement would also help each teacher education institution to more clearly define its course outcomes and other related educational experiences. Finally, instead of having state licensure and certification exams taken upon completion of a program and that essentially measure only one competency area (i.e., content knowledge), exams measuring each competency area could be taken upon completion of a defined number of prerequisite courses.

While these competency area exams would reference course instruction and outcomes, they would be most closely linked to the state model of teaching competency (i.e., the state certification model) and require that students demonstrate particular competencies to an examiner. These assessments may include ratings, oral presentations, formal examinations, simulated classroom situations, and any other number of measuring techniques directly linked to the requirements and outcomes of the courses designed to produce the competency assessed. Figure 4 presents an example of the kinds of items that might be developed to measure

Insert Figure 4 About Here

the competency areas "Classroom Instruction" and "Adaptability." Since the content of these competency area assessments would measure the outcomes for a particular series of courses linked to one or more of the competency areas, they provide the empirical link between a particular Teacher Education program and the actual competencies required to be a successful teacher.

4. A list of all of the resources needed to implement instruction, including: laboratory and public school curriculum materials, videos, filmstrips, lab sheets, study guides, test banks, handouts, required readings, and overhead transparencies; and
5. A list of the course activities and requirements, including a detailed description of the criteria for the successful completion of each.

V. **A Program Technical Report For Each Specialty Area Program**

The purpose of the Program Technical Report is to fully document the analytical framework and content validity of each specialty area program. It describes and documents each of the previously listed program products, as well as any summative evaluation information regarding the specific costs and benefits of the program. Its function is to provide the necessary program information required by certification and accreditation groups. It should be authored by the key faculty member(s) in that program; however, like

the Course Guide, there should be a standard format for these reports that includes:

1. The rationale for the program;
2. The programs's goals, objectives, outcomes, and teaching competency and knowledge base model;
3. A brief description of each course of instruction within the program (this information could be taken from the Course Guide);
4. A list of the teaching competencies associated with each course or group of courses (the matrix produced in Number III would suffice);
5. The test blueprint for each of the competency area assessment instruments designed to operationally measure the program's teaching competency areas;
6. Any summative evaluation information that might exist on the program's costs and benefits.

VI. A Set of Analytically Defined Direct Costs Required For Training a Teacher

The question that this product is designed to answer is, "How much does it cost to train a teacher at this institution?" Therefore, estimates of all direct costs required to train a teacher should be compiled

and reported. These costs should include estimates of all investments in time and resources, including personnel salaries, buildings, and other facilities. Other costs would include federal, state, and local allotments to the program and student grant, loan, and scholarship funds.

VII. Longitudinal Student Data Base

An extensive, but succinct, longitudinal base is a central component of an effective teacher education program assessment system. Its purpose is to track and monitor students as they enter the institution, progress through their course work, successfully complete the competency assessments, and graduate. It also tracks their subsequent experience as an alumni. It is compiled through an ongoing data collection and assessment system carried out through the institution's assessment office by the university examiner, institutional researcher, or the like. This data base enables administration and faculty to accomplish a number of relevant objectives, including:

- Predicting the academic success of students who enroll in order to: a) help ensure that a high

percentage of those who enroll will graduate and
 b) capitalize upon this information for advisement
 purposes;

- Developing placement procedures to support various
 specialty area programs;
- Linking programs, courses, and other kinds of
 college experiences to specific educational
 outcomes and teaching competencies to make
 empirical, cost-benefit decisions in regard to
 programs, courses of instruction, and other
 degree-related requirements; and
- Improving programs and courses of instruction.

Creating, refining, and maintaining this data base
 requires establishing a carefully designed data
 collection system. The purpose of this data collection
 system is to obtain and maintain the achievement and
 biosocial data required for the summative and
 confirmative phases of the evaluation.

To create the Student Data Base requires first,
 determining the specific kinds of data that should be
 included in the data base; and, secondly, developing or
 identifying measures of these data. Appendix D
 presents an array of examples of the kinds of

information that could possibly be compiled on Teacher Education students and graduates (See also, Nichols, 1989). To collect some of this information will require the development or identification of additional student assessment measures. These measures are discussed in the summative evaluation section of this paper.

VIII. A Design For a Summative Evaluation

The final task in the formative evaluation phase is to identify a design for one or more summative evaluations. Given the results of the formative evaluation, the summative evaluation phase can then evaluate the program's effects in terms of its stated goals, outcomes, and competencies and compare these anticipated ones with the empirical ones. In this way a "program-fair" evaluation can be carried out (Hanson, Schutz, and Bailey, 1981).

Phase II: Summative Evaluation

Summative evaluations are carried out on each specialty area program for several reasons. Most important is to ensure that the program is being implemented as designed and to measure the degree or level of the implementation. Other reasons are to improve the function of the program, enhance program implementation, and provide measures of preservice teacher competency levels. Moreover, the immediate direct costs and effects of each program can be firmly established.

As Figure 5 indicates, in summative evaluation, the monitoring of program implementation takes place

 Insert Figure 5 About Here

through on-going information-gathering. The products developed and the activities that take place during this phase focus on assuring that the program is being fully implemented and that the course outcomes and program competencies are being attained. In essence, it establishes the concurrent validity of the operational program.

The products that are produced during the summative evaluation phase include: implementation

requirements, procedures and measures; on-going reports on implementation; summary reports on program costs and effects; and student entry and exit assessments. In addition, a program orientation session should also be developed to familiarize new faculty members with the program and course architecture and course implementation procedures.

Implementation procedures are developed to ensure that the teaching staff can fulfill all required procedures to conduct instruction, administer exams, and report results. Thus, measures of these procedures must then be either developed or identified for the purpose of indicating the level of course implementation practices. Students' scores on each of these implementation measures (e.g., tests, reports, projects) are included in the Student Data Base to: a) document program implementation practices, b) document student competencies, and c) validate immediate program effects. The program reports are generated by the institution's assessment office for use by administration and management.

Once again, no precise time period has been established for the completion of the summative

evaluation phase; however, it is estimated to take place over a 3- to 5-year period. The criteria for the completion of this phase are provided in Appendix A.

Summative Evaluation Products to be Developed

I. Program Orientation Session

The purpose of this session is to: a) familiarize faculty members with the results of the formative evaluation efforts and products, b) provide them with the rationale for and an overview of the summative evaluation efforts, and 3) discuss with them their expected contribution to these efforts. This session would be held, either on a yearly basis or as needed, as part of the orientation process for new faculty members or as staff development for incumbent faculty.

II. Summative Evaluation Program Reports

This series of reports should document:

1. the relationships that exist between the number of courses taken in a specialty area program and the level of teacher competencies;

2. the course implementation practices of instructors (i.e., are the instructors providing instruction in and requiring students to practice the course's defined objectives and requirements?);
3. whether or not the program's competency areas are being attained by students, independent of the background characteristics of either the instructors or the students.
4. the projected costs of each specialty area program for training a teacher in that program;
5. the context and implementation factors that impact both program costs and outcomes;
6. the concurrent validity of each specialty area program (i.e., the course requirements (implementation measures) are linked to program outcomes (course grades) and program outcomes are linked to program competency areas (competency area assessments) and teacher certification).

Collectively, the information compiled in these reports allows administrators to make intelligent responses,

based on empirical rather than analytical or speculative data, when they are asked to eliminate, reconfigure, or reduce funding for their programs or components of their programs.

III. Student Entrance and Exit Assessment Instruments

To collect all of the information needed for the summative and confirmative phases of the evaluation requires the development of at least two separate student assessment instruments:

1. a Student Entrance Assessment instrument; and
2. a Student Exit Assessment instrument.

The Student Entrance Assessment would assess both the bio-social and the achievement background of students and provide measures of a variety of characteristics relevant to success in both higher education in general and in a teacher education program. Although some of the measures identified for use in this assessment (e.g., standardized tests of math, English grammar, reading) could be used for placement purposes, placement tests should eventually be developed for all courses within a specialty area program, including the general education component. These comprehensive CLEP

assessments, which might feasibly take place over the course of several days, would be given to students, upon their request, to provide the means for more competent students to graduate early.

The Student Exit Assessment instrument would probably be much less extensive than the entrance assessment and would include measures of short-term educational outcomes, such as measures of achievement, attitudes/behaviors, and competence in particular predefined areas. Ideally, many of those students who decide to drop-out before graduation would also complete this measure so that naturally occurring comparison groups could be formed.

The responsibility for administering this assessment to graduating seniors would be student services; perhaps financial aid or admissions could take the responsibility for administering it to those students who drop-out before completing their degree requirements. In other words, completing this assessment should be considered part of any required exit procedure, graduating or not. This would also involve mailing the assessment instrument to some students and anticipating getting them back.

It should be noted that the systematic development of these instruments is absolutely vital to the effort to link student experiences to educational outcomes. An initial step in this process would be to design the student assessment instruments to specifically measure a program's defined outcomes and competency areas (see Hanson and Siegel, 1988a). Furthermore, each of the student assessment instruments must be researched based (i.e., all items included in the instruments must be either empirically or logically related to the predefined educational outcomes, field tested, and validated). Appendix D gives examples of possible items that could be include in each instrument. Once completed, scores would be obtained on the computed variables measured in these assessments and include in the Student Data Base.

IV. A Defined Set of Course Implementation Measures

For Each Course Of Instruction

For each course of instruction three or more milestone assessments must be identified or developed as a means of measuring the level of course implementation practices. These assessments could be

papers, projects, tests, or any other assignment that would indicate that the instructor covered the material and provided students with instruction and practice in the outcome skills designated in the Course Guide. The students' grades on these assessments would then be entered into the Student Data Base to indicate both the level of the student's skill in this outcome area and the level of course implementation.

Phase III: Confirmative Evaluation

The final phase of the certification evaluation model, confirmative evaluation, is defined and summarized in Figure 6. Unlike the previous two phases, the literature does not use this term;

Insert Figure 6 About Here

nor do schools and evaluators usually consider the issue it refers to: the enduring benefits of a program. Enduring or long-term benefits refers to those changes that can be identified only after the passage of time and are directly linked to participation in a specific program (e.g., Novack and Musonda, 1991; Hanson and Siegel, 1988b;1993)

program's identified teacher competency areas to successful careers in teaching. Examples of the kinds of items to be included in this instrument are provided in Appendix D. The information from this assessment is then fed back into the program review process to adapt, delete, or add particular components that, as a result of the summative evaluation, have been empirically identified as necessary to the model. The Criteria for completion of a confirmative evaluation are contained in Appendix A.

Confirmative Evaluation Products to be Developed

I. Confirmative Evaluation Program Reports

The information generated in these reports serves to improve implementation practices in the specialty area programs and to link their course outcomes to program model competencies. Answers can now be provided to such question as:

- Does this program lead most individuals to achieve a successful teaching career?
- What competencies should be added, deleted, or adapted to improve the model?

- Are the individuals completing this program able to successfully adapt to new or different teaching situations?
- Are the individuals who complete this specialty area program more successful teachers than those who graduated in a content area program and received an alternative teaching certificate?
- What are the specific patterns of educational experiences that show a consistent relationship to successful teaching careers?
- Is this program cost-beneficial over the long-term and/or is there still a need for this program?

II. The Student Follow-up Assessment Instrument

This instrument would be mailed to students every 2 to 5 years after leaving the college or university and provide measures of long-term educational benefits. This instrument would confirm the enduring effects of particular courses, programs, and other relevant college experiences. This kind of evaluation produces the real "pay-off" for Teacher Education Programs. With this kind of information, administrators and other interested audiences can determine what "really counts"

in terms of identifying the specific kinds of experiences necessary to produce competent and responsible adults (see Siegel, 1990). That is, administrators can base their decisions to revise or delete particular programs, courses and/or other degree requirements based on empirical data. Those assignments and requirements identified as either having no impact or producing negative outcomes can be minimized or eliminated and those showing a positive effect can be refined and enhanced. More importantly, more ideal and effective programs, specifically tailored to benefit the needs of students and society, will continue to evolve over time.

One logical office to take responsibility for coordinating this assessment process would be the Alumni Office in cooperation with the Alumni Association. It would also be anticipated that the Alumni Association would have considerable input as to the kinds of items they would like to include in this assessment. Data obtained from this measure would be extremely helpful to them in planning alumni activities, for developing newsletters and yearbooks, and presenting awards.

Conclusion

As it stands today, the vast majority of Teacher Education Programs implemented in our nation's colleges and universities are best described as an analytically configured set of courses, activities, and assessments compiled by politicians and educators who "assume" the configuration will achieve the desired objectives, produce the intended outcomes, and develop the desired competencies. Few programs, if any, are designed or documented in such an operational manner that their outcomes can be empirically validated. As a result, huge amounts of money are spent every year on "Programs" that have no established validity; hence, the results of any evaluation efforts undertaken on them do not provide any useful information.

This paper briefly outlines, in operational terms, the components and products of a longitudinal evaluation model for teacher education programs. While virtually every teacher education program, and all educational expenses are assumed to have some enduring impact, efforts to actually identify and confirm these benefits are few. With the need for accountability and the presence of a better information systems, the

incentives and mechanisms are all in place. Moreover, the tasks outlined in this paper set the agenda for a complete Teacher Education Program evaluation system that would track policies and students in each specialty area program during its operational life. In this way, firm estimates of a program's long-term effects/benefits and costs can be derived.

References

- Cooley, W. (1983). Improving performance of an educational system. Educational Researcher, 12(6), 4-12.
- Cronbach, L. (1982). Designing evaluations of educational and social programs. San Francisco, CA: Jossey-Bass.
- Cronbach, L. (1980). Toward reform of program evaluation. San Francisco: Jossey-Bass Publishers.
- Dingan, M.B. (1986). Measurement and evaluation of health education. Springfield, IL: Charles Thomas Publishers.
- Hanson, R.A. (1986). Program evaluation and improvement based on cost-effectiveness analysis. Studies in Educational Evaluation, 12, 89-108.
- Hanson, R., Lehnman, J., & Bailey, J. (1981). A six-year longitudinal study of an instructional innovation. Paper presented at the 1981 annual meeting of the American Educational Research Association, Los Angeles, CA.
- Hanson, R., McMorris, R. & Bailey, J. (1986). Differences in instructional sensitivity between item formats and between achievement test items. Journal of Educational Measurement, 23(1), 1-12.
- Hanson, R. & Schutz, R. (1978). A new look at schooling effects from Programmatic Research and Development. In D. Mann (Ed.), Making Change Happen? New York: Columbia University Press.
- Hanson, R., Schutz, R., & Bailey, J. (1977). Program-fair evaluation of instructional programs: Initial results of the kindergarten reading readiness inquiry (Technical Report No. 57). Los Alamitos, CA: SWRL Educational R&D.

- Hanson, R. & Schutz, R. (1981). Rethinking evaluation design for nutrition education. Journal of Nutrition Education, 13, 86-90.
- Hanson, R. & Siegel, D (1989, April). Cost-benefit analysis of kindergarten reading instruction. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Hanson, R. & Siegel, D.F. (1988b). The long-term effects on high school seniors of learning to read in kindergarten. Technical Report #1. Garden Grove, CA: Hanson Research Systems. ERIC Document Reproduction Service No. ED 324 128.
- Hanson, R. & Siegel, D.F. (1988a). Design and development of The Reading Biographer: An application of function specific assessment. Paper presented at the annual meeting of the National Council for Measurement in Education. ERIC document 324 355.
- Hanson, R.A. & Siegel, D.F. (1993). Assessing the enduring effects of the Energy Source Education Program. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Hanson, R.A. & Siegel, D.F. (1991). Formative, summative, and confirmative: The three phases of evaluation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Haywood, J. (1989). Assessment in higher education. New York: Wiley
- Jordan, E.W. & Machesky, J.J. (1990). Systems development: Requirements, evaluation, design, and implementation. Boston, MA: PWS-Kent publishing Co.
- Kelly, D. (1994, January 4). More take alternative path to teaching. USA Today, p.1D

- Kinney, L.B. (1964). Certification in Education. Englewood Cliffs, NJ: Prentice-Hall.
- Nichols, J.O.(1989). Institutional effectiveness and outcomes assessment implementation on campus: A practitioner's handbook. New York: Agathon Press.
- Novak; J.D. & Musonda, d. (1991). A twelve-year longitudinal study of concept learning. American Educational Research Journal, 28, 117-153.
- Pearce, R.A. (1991, June, 18). Major changes in higher education. Tulsa Tribune., 97,(141),p.1.
- Provus, M.M. (1971). Discrepancy Evaluation. Berkeley, CA: McCutchan Publishing.
- Raizen, S.A. & Rossi, P.H.(Eds.)(1981). Program evaluation in education: When? how? to what ends? Washington, D.C.: National Academy Press.
- Raschke, C. (Ed.). (1994). Education Trend Letter, 1(3).
- Siegel, D. & Hanson, R.(1991). The three phases of evaluation: formative, summative, and confirmative. OERS IV Conference Proceedings. Norman, OK.
- Siegel, D.F.(1990). The literacy press: A process model for reading development. The Journal of Educational Research, 83(6),336-345.

Program Stages and Activities	Time in Years	Evaluation Phase
<u>I. Design</u>		
Defining the goals, objectives content, potential component activities and projected outcomes.	1-2	Formative
<u>II. Development</u>		
Creating all the resources and procedures that embody a program and that are necessary to implement it.	2-3	Formative
<u>III. Implementation/Dissemination</u>		
The resources needed to monitor program implementation and outcomes in diverse settings and populations.	1-4	Formative and Summative
<u>IV. Monitored Use and Management</u>		
Operation of a system to validate and assure program outcomes and that program effects and costs are maintained.	5-10+	Summative and Confirmative

Figure 1. The relationship among the program stages, expected time periods, and evaluation phases

Definition of Formative Evaluation

FORMATIVE EVALUATION: Gathering information during program development to modify and improve the foundation of an educational program until it can accomplish its defined goals.

- WHEN:** Ideally, during design/development and before publication and wide-scale use; Otherwise, before any further evaluations take place.
- PURPOSE:** To define its intentions and build a program or product that will accomplish them (i.e., to validate the program/product, procedures, and resources embodied in a program in the context of actual use with small groups to achieve defined conditions/ specifications).
- FOCUS:** Program design, development, modification, and improvement to enhance content validity.
- BY WHAT MEANS:** Analysis of resources and intentions. Studies of specific components of procedures with small groups or professional program reviews.
- WITH WHOM:** Small groups of users, internal and external consultant/analysts.
- BY WHOM:** Program developers (e.g., authors, administrators, instructors)
- CRITERIA FOR COMPLETION:**
1. A set of costs and benefits, consistent with the program's goals, are analytically defined in a technical report.
 2. The program consists of a complete set of resources and procedures that have been validated in controlled settings.
- EXAMPLE:** Hanson, R.A. & Schutz, R.E. (1981). Rethinking for nutrition education. Journal of Nutrition Education, 13, 86-90.
-

Figure 2. Definition of formative evaluation.

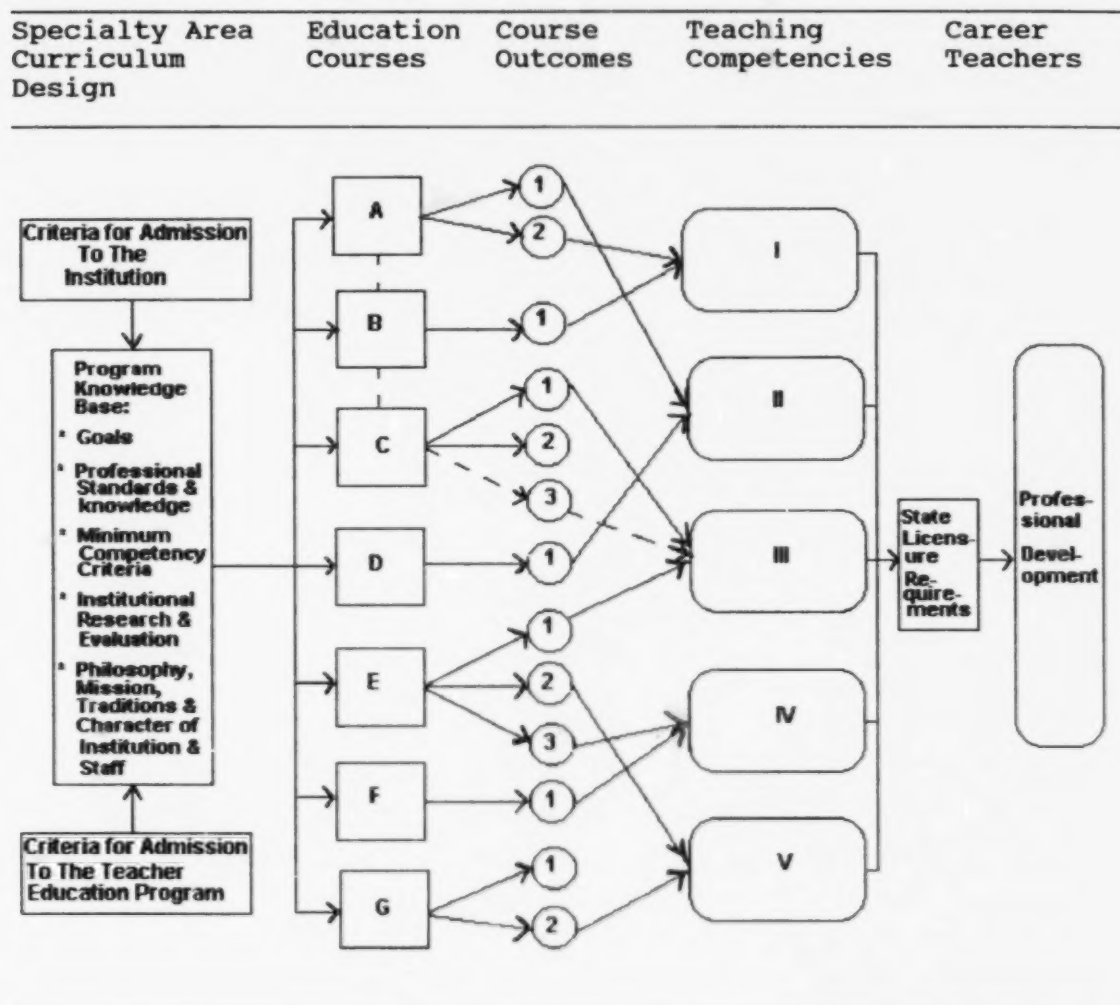


Figure 3. Teacher education specialty area program competency development model.

Competencies Measured: 1. Classroom Instruction
 2. Adaptability

Task: Adapt instructional resources to develop an aligned set of lesson plans designed to teach specific skills and/or concepts appropriate for a specified elementary grade level.

Performance Item

Given a teacher's resource kit for a particular subject area, the student will develop a set of aligned lessons designed to teach a specific set of skills and/or concepts. The lessons must cover a period of two weeks, be developed around the 5-part lesson plan format, and include an assessment component.

Multiple choice Item

The components of the five-part lesson plan format include all of the following EXCEPT:

- A. Independent Practice
- B. Learner Outcomes
- C. Guided Practice
- D. Introduction
- E. Content Presentation

Interpretive Items

Given lessons taken from a teacher's guide for different elementary subject areas, a series of multiple choice questions will be asked about each lesson's alignment and appropriateness of the enabling activities for students of a particular grade level and with particular background characteristics.

Figure 4: Example of competency area assessment items

Definition of Summative Evaluation

SUMMATIVE EVALUATION: Gathering information during the operational use of a program to determine the direct costs, benefits, and the context and implementation factors that impact them.

WHEN:	During the operational life of a program.
PURPOSE:	To determine what direct effects occur and under what conditions (i.e., costs and benefits) and to manage/monitor implementation.
FOCUS:	All components/units/segments of programs as they are used in operation.
WITH WHOM:	All users, if possible.
BY WHOM:	University assessment offices, districts, outside evaluators using the resources provided by publishers/developers, and others as directed by an analysis of the program.
BY WHAT MEANS:	Instrumentation and computer support from institution/developer as required by evaluators, administrators, and others involved.
CRITERIA FOR COMPLETION:	1. Empirical estimates of direct effects and costs; 2. Established relationships between time & use and use & effects. 3. Complete implementation resources and monitoring system.

(figure continues)

Figure 5 Cont.

Definition of Summative Evaluation

EXAMPLES &
FURTHER
DISCUSSION:

Hanson and Siegel (1989, April).
 Cost-benefit analysis of
kindergarten reading
instruction. Paper presented at
 the 1989 annual meeting to the
 American Educational Research
 Association, San Francisco, CA.

Hanson, R.A. & Schutz, R.E. (1978). A
 new look at schooling effects from
 programmatic research and
 development. In D. Mann (Ed.),
Making change happen? New York:
 Columbia University Press.

Hanson, R.A., Schutz, R.E. & Bailey, J.
 (1977). Program-fair
evaluation of instructional
programs: Initial results of the kindergarten
reading readiness inquiry (Technical Report
 No. 57). Los Alamitos, CA: SWRL Educational
 R&D.

Figure 5. Definition of summative evaluation.

Definition of Confirmative Evaluation

CONFIRMATIVE EVALUATION: Gathering information to determine the enduring effects and benefits associated with an educational program.

WHEN:	At specified intervals following use (minimum interval 2 years).
PURPOSE:	To determine the enduring effects and long-term benefits of the program.
FOCUS:	Configurations of program use based on records of implementation.
WITH WHOM:	Samples of users, non-users, and other comparison groups of interest based on biosocial, political, and pedagogical concerns of schools and districts.
BY WHOM:	State departments of education, user consortiums, outside consultants and evaluators, large districts, and institutions of higher education.
BY WHAT MEANS:	Detailed school records and follow-up studies using instrumentation developed to measure program effects ranging from those that are unique to the program to those that are common to the School/community.
CRITERIA FOR COMPLETION:	1. Program is linked to broader accomplishments (i.e., board direct effects); 2. Actual costs and benefits are established; 3. Outcomes are confirmed.
EXAMPLES & FURTHER DISCUSSION:	Hanson and Siegel (1988). The long-term <u>effects of learning to read in kindergarten</u> (Technical Report No.1). Garden Grove, CA: Hanson Research Systems.

Figure 6. Definition of confirmative evaluation.

APPENDIX A
Criteria for Completion of A
Formative, Summative, and Confirmative Evaluation

Criteria for the Completion of
A Formative Evaluation

1. The program's target population is identified.
2. There is a general statement of the program's goal that, at least, specifies what the program seeks to accomplish, with whom, and in what settings (e.g., the goal of this program is to provide training in effective teaching in the elementary school and enable individuals to become certified in Elementary Education).
3. Production of a preliminary model of teacher competency. This consists of a set of competencies most directly related to the course outcome areas. For each competency, minimum performance standards should be indicated.
4. There is a set of plausible behavioral objectives and outcomes for each course of instruction that have been analytically derived and that are logically linked to: a) one another, b) the identified teacher competencies for the program and d) the overall goal for the program.
5. The specific courses, and the skills and concepts taught within each course in the program, are identified and the criteria for completing each requirement are clearly specified (i.e., the course requirements for meeting the objectives and producing the outcomes for each course within a specialty area program).
6. There is a complete, operational set of analytically derived resources, procedures, and activities that: a) embody the instruction designed to produce the learning for each objective; b) logically would be anticipated to lead to the desired outcome(s); and c) can be replicated in other classrooms.
7. There are appropriate assessment instruments which can realistically hold the program accountable (i.e., measures of the implementation of course requirements as well as measures of the program's outcomes).

(criteria continues)

criteria continued

Criteria for the Completion of
A Formative Evaluation

8. There is a preliminary estimate of costs and benefits for each course that are consistent with the program goals that have been analytically defined.
 9. The likely uses of the program's performance information have been identified.
 10. A design for one or more summative evaluations has been formulated.
-

Criteria For Completion of A Summative Evaluation

1. Student entrance and placement assessments have been fully developed for the purpose of properly placing students in prerequisite, beginning, or advanced courses and obtaining the necessary bio-social background information on students to carry out effective longitudinal program evaluations.
 2. Measures of program implementation are in use and included in the Student Data Base (i.e., measures of course activities and other requirements students must successfully completed for each course of instruction).
 3. Instructionally sensitive outcome assessments are in use and included in the Student Data Base (i.e., assessments/tests designed to measured the defined outcomes for the program and course(s) of instruction).
 4. A Student Exit Assessment have been fully developed, is in use, and included in the Student Data Base.
 5. The Student Data Base is on-line, monitored, and updated.
 6. A program orientation and implementation session for new faculty members has been developed and conducted.
 7. Program concurrent validity and cost have been established and reported in the summative evaluation reports.
-

Criteria for Completing a Confirmative Evaluation

1. A follow-up assessment instrument, which covers professional experiences, prior educational experiences, and future plans, has been systematically designed, developed, and placed in use and the information collected is included in the Student Data Base.
 2. Three or more years of successful program operation documented by summative evaluation data on measured course implementation and completion, professional teaching competencies, and related educational experiences.
 3. The Student Data Base is fully developed and contains longitudinal data on all of the above information.
 4. Personnel and other resources are permanently in place to effectively maintain the evaluation system.
-

APPENDIX B

Examples of Professional Teaching Competencies
and Indicators of Each Competency

Examples of Professional Teaching Competencies and
Indicators of Each Competency

Professional Teaching Competency	Indicators of Professional Teaching Competency	
Classroom Organization and Management	Demonstrates effective and appropriate organization and management of:	
	1. Time, Resources, & Classroom	5. Grading Patterns/Policy
	2. Lesson Plans	6. On-task Behavior/Orderly Environment
	3. Student Files/Records	7. Within and Between Classroom Transitions
	4. Classroom Rules/Discipline Plan	
Classroom Instruction	Demonstrates effective and appropriate:	
	1. Use of Resources	6. Content Coverage
	2. Instructional Objectives	7. Student Achievement
	3. Questioning Techniques	5. Lesson Sequencing/Presentation Format
	4. Clarity of Instruction	9. Monitoring and Feedback
	5. Instructional Methods	10. Motivation and Praise
Interpersonal Competency and Human Relations	Demonstrates appropriate:	
	1. Speech and Grammar	5. Communication Ability
	2. Dress and Personal Appearance	6. Tact and Courtesy
	3. Initiative and Cooperation	7. Behavior & Composure
	4. Interaction with Administration and Staff/Parents and Students	
Adaptability	Demonstrates successful adaptability in:	
	1. Lesson Plans and Resources	4. Use of New Technology
	2. Related Grade Level Assignments	5. Use of New Ideas or Ideas of Others
	3. Related Subject Area Assignments	6. Changing Educational Policies and Practices
Professionalism	Maintains involvement in:	
	1. Additional Course Work	4. Community or School Projects and Functions
	2. Workshops/In-Service	5. Professional Organizations
	3. Research/Publications	
Assessment and Critical Thinking	Demonstrates effectiveness and appropriateness in:	
	1. Problem Solving	3. Evaluation of Pupil Performance
	2. Decision Making	4. Organization
Content Area Knowledge and Skills	Demonstrates appropriate background knowledge and skills in:	
	1. Subject Areas Taught	3. Liberal Arts and Science Areas
	2. Health and Safety Areas	4. Educational Legal and Ethical Areas

APPENDIX C

Information for Possible Consideration
in a Longitudinal Student Data Base

Information for Possible Consideration
in a Longitudinal Student Data Base

1. ID (Social security number)
2. Biosocial/demographic Data (name, home address, school address, birth date, gender, racial/ethnic background, date of entry into institution)
3. SAT/ACT Scores
4. High School G.P.A.
5. Declared Major(Specialty Area Program)
6. Entrance Assessment scores
7. Course Implementation Scores
8. English Competency Exam score
9. PPST/or other Test scores used for admission to the Teacher Education Program
10. Date of admission to Teacher Education
11. Score on each concept/skill/competency area assessed
12. Membership in Honor Societies
13. Academic Probation
14. Location of field, clinical, and student teaching experiences
15. Curriculum Exam score(s)
16. College Graduation G.P.A.
17. Exit Exam Scores
18. Date of Teacher Licensing
19. Follow-up Assessment scores

Appendix D

Examples of Items to Include in the
Entrance, Exit, and Follow-up Assessment Instruments

Example of Possible Items to Include in the
Entrance Assessment Instrument

Part I: Personal Background Information

1. High school academic, family, and social background

A. measures of academic achievement, e.g.,

1. type of high school attended (rural, suburban, or urban)
2. High School diploma or GED
3. Overall high school G.P.A.
4. SAT/ACT
5. High school honors/remedial classes
6. High school academic track
7. Number of college semesters completed/associate degree

B. Measures of family background, e.g.,

1. income level of family
2. educational level of father/mother/guardians
3. occupation of father/mother/guardians
4. parent's educational expectations

C. Measures of participation in discretionary activities, e.g.,

1. extracurricular school activities
2. work experiences
3. hobbies

4. other kinds of community organizations
2. College academic background, e.g.,
 1. Overall college G.P.A
 2. College level remedial classes
 3. College honor societies
3. Personal Independence
 - A. measures of independence from family ,e.g,
 1. number of marriages
 2. number of children
 3. full-time job
 4. care for parent or older relative
 - B. measures of financial independence, e.g.,
 1. parents contribution to expenses
 2. full scholarship/grant
 3. grants/loans
 4. student employment
 5. level of indebtedness
 - C. Measures of medical and legal independence, e.g,
 1. overall condition of physical health
 2. overall condition of mental health
 3. arrests, convection, or other encounters with the criminal justice system
 4. current legal or medical entanglements

4. Motivation/Attitudes/Self-Esteem

- A. measures of how motivated the student is to complete degree requirements (i.e., a motivational scale);
- B. Attitudes toward social responsibility, esthetics and level of self-esteem (i.e., and Attitude scale).

Part II: Standardized tests of Reading and Vocabulary
Reading and Vocabulary Achievement (e.g., ACT/SAT Scores).

Part III: Standardized tests of Math and Science
Math and Science Achievement.

Examples of Possible Items to Include in the
Exit Assessment Instrument

Measures of short-term benefits

- A. baccalaureate degree
- B. college level reading ability
- C. job opportunities/interviews
- D. overall intellectual competency
 - 1. communications
 - 2. analytical and problem solving
 - 3. social responsibility
 - 4. esthetics awareness
- E. change in self-esteem, general outlook on life
- F. change in attitudes and behaviors
- G. years in college
- H. Overall G.P.A
- I. Major G.P.A.

Examples of Possible Items to Include in the
Follow-up Assessment Instrument

- A. Biographical information (i.e., address, age, marital status, number of children)
- B. Further evaluation of educational experiences and courses taken at the institution
- C. Current job status
- D. Current level of income
- E. Number of different teaching jobs
- F. Awards/achievements
- G. Certification/post hoc educational experiences
- H. Level of personal satisfaction e.g., home, family, self-esteem
- I. Level of job satisfaction
- J. Current job opportunities available
- K. Travel/recreational activities
- L. General outlook on life
- M. Involvement in community/church organizations and activities
- N. Involvement in professional organizations/activities
- O. Future goals
- P. General condition of health